



# SEQUENCE LISTING

## (1) GENERAL INFORMATION:

(i) APPLICANT: O'Malley, Bert W.  
Tsai, Ming-Jer  
Ledebur, Harry C. Jr.  
Kittle, Joseph D. Jr.

(ii) TITLE OF INVENTION: MODIFIED STEROID  
HORMONES FOR GENE  
THERAPY AND METHODS  
FOR THEIR USE

(iii) NUMBER OF SEQUENCES: 14

## (iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Lyon & Lyon  
(B) STREET: 633 West Fifth Street  
Suite 4700  
(C) CITY: Los Angeles  
(D) STATE: California  
(E) COUNTRY: U.S.A.  
(F) ZIP: 90071-2066

## (v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
storage  
(B) COMPUTER: IBM Compatible  
(C) OPERATING SYSTEM: IBM P.C. DOS 5.0  
(D) SOFTWARE: Word Perfect 5.1

## (vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: 08/959,013  
(B) FILING DATE: October 28, 1997  
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

- (A) APPLICATION NUMBER:  
(B) FILING DATE:

(viii) ATTORNEY/AGENT INFORMATION:

- (A) NAME: Warburg, Richard J.  
(B) REGISTRATION NUMBER: 32,327  
(C) REFERENCE/DOCKET NUMBER: 226/286

(ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: (213) 489-1600  
(B) TELEFAX: (213) 955-0440  
(C) TELEX: 67-3510

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6177 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

CTAGAGTCGA	CCTGCAGCCC	AAGCTCTCGA	GGGATCCTGA	GAAC TTCAGG	GTGAGTTTGG	60
GGACCCTTGA	TTGTTCTTTC	TTTTTCGCTA	TTGTAAAATT	CATGTTATAT	GGAGGGGGCA	120
AAGTTTTCAG	GGTGTGTGTT	AGAATGGGAA	GATGTCCCTT	GTATCACCAT	GGACCCTCAT	180
GATAATTTTG	TTTCTTTCAC	TTTCTACTCT	GTTGACAACC	ATTGTCTCCT	CTTATTTTCT	240
TTTCATTTTC	TGTAAC TTTT	TCGTTAAACT	TTAGCTTGCA	TTGTAAACGA	ATTTTAAAT	300
TCACTTTTGT	TTATTTGTCA	GATTGTAAGT	ACTTCTCTA	ATCACTTTTT	TTTCAAGGCA	360
ATCAGGGTAT	ATTATATTGT	ACTTCAGCAC	AGTTTTAGAG	AACAATTGTT	ATAATTAAAT	420
GATAAGGTAG	AATATTCTG	CATATAAATT	CTGGCTGGCG	TGGAAATATT	CTTATTGGTA	480
GAAACAATA	CATCTGGTC	ATCATCCTGC	CTTCTCTTT	ATGGTTACAA	TGATATACAC	540
TGTTTGAGAT	GAGGATAAAA	TACTCTGAGT	CCAAACCGGG	CCCCTCTGCT	AACCATGTTC	600
ATGCCTTCTT	CTTTTTCCTA	CAGCTCCTGG	GCAACGTGCT	GGTTGTTGTG	CTGTCTCATC	660
ATTTTGGCAA	AGAATTCACT	CCTCAGGTGC	AGGCTGCCTA	TCAGAAGGTG	GTGGCTGGTG	720
TGGCCAATGC	CCTGGCTCAC	AAATACCACT	GAGATCTTTT	TCCCTCTGCC	AAAAATTATG	780
GGGACATCAT	GAAGCCCTT	GAGCATCTGA	CTTCTGGCTA	ATAAAGGAAA	TTTATTTTCA	840

TTGCAATAGT	GTGTTGGAAT	TTTTTGTGTC	TCTCACTCGG	AAGGACATAT	GGGAGGGCAA	900
ATCATTTAAA	ACATCAGAA	GAGTATTTGG	TTTAGAGTTT	GGCAACATAT	GCCATATGCT	960
GGCTGCCATG	AACAAAGGTG	GCTATAAAGA	GGTCATCAGT	ATATGAAACA	GCCCCCTGCT	1020
GTCCATTCC	TATTCATAG	AAAAGCCTTG	ACTTGAGGTT	AGATTTTTTT	TATATTTTGT	1080
TTTGTGTTAT	TTTTTCTTT	AACATCCCTA	AAATTTTCCT	TACATGTTTT	ACTAGCCAGA	1140
TTTTTCCTCC	TCTCCTGACT	ACTCCAGTC	ATAGCTGTCC	CTCTTCTCTT	ATGAACTCGA	1200
GGAGCTTTTT	GCAAAAGCCT	AGGCCTCCAA	AAAAGCCTCC	TCACTACTTC	TGGAATAGCT	1260
CAGAGGCCGA	GGCGGCCTCG	GCCTCTGCAT	AAATAAAAAA	AATTAGTCAG	CCATGGGGCG	1320
GAGAATGGGC	GGAACCTGGC	GGAGTTAGGG	GCGGGATGGG	CGGAGTTAGG	GGCGGGACTA	1380
TGGTTGCTGA	CTAATTGAGA	CTGCATTAAT	GAATCGGCCA	ACGCGCGGGG	AGAGGCGGTT	1440
TGCGTATTGG	GCGCTCTTCC	GCTTCCTCGC	TCACTGACTC	GCTGCGCTCG	GTCGTTCCGG	1500
TGCGGCGAGC	GGTATCAGCT	CACTCAAAGG	CGGTAAATACG	GTTATCCACA	GAATCAGGGG	1560
ATAACGCAGG	AAAGAACATG	TGAGCAAAAG	GCCAGCAAAA	GGCCAGGAAC	CGTAAAAAAGG	1620
CCGCGTTGCT	GGCGTTTTTC	CATAGGCTCC	GCCCCCTGA	CGAGCATCAC	AAAAATCGAC	1680
GCTCAAGTCA	GAGGTGGCGA	AACCCGACAG	GACTATAAAG	ATACCAGGCG	TTTCCCCCTG	1740
GAAGCTCCCT	CGTGCGCTCT	CCTGTCCGA	CCCTGCCGCT	TACCGGATAC	CTGTCCGCCT	1800
TTCTCCCTTC	GGGAAGCGTG	GCGCTTTCTC	AATGCTCACG	CTGTAGGTAT	CTCAGTTCCG	1860
TGTAGGTCGT	TCGCTCCAAG	CTGGGCTGTG	TGCACGAACC	CCCCGTTTCA	CCCGACCGCT	1920
GCGCCTTATC	CGGTAACAT	CGTCTTGAGT	CCAACCCGGT	AAGACACGAC	TTATCGCCAC	1980
TGGCAGCAGC	CACTGGTAAC	AGGATTAGCA	GAGCGAGGTA	TGTAGGCGGT	GCTACAGAGT	2040
TCTTGAAGTG	GTGGCCTAAC	TACGGCTACA	CTAGAAGGAC	AGTATTTGGT	ATCTGCGCTC	2100
TGCTGAAGCC	AGTTACCTTC	GGAAAAAGAG	TTGGTAGCTC	TTGATCCGGC	AAACAAACCA	2160
CCGCTGGTAG	CGGTGGTTTT	TTTGTTTGCA	AGCAGCAGAT	TACGCGCAGA	AAAAAAGGAT	2220
CTCAAGAAGA	TCCTTTGATC	TTTTCTACGG	GGTCTGACGC	TCAGTGGAAC	GAAAACTCAC	2280
GTTAAGGGAT	TTTGGTCATG	AGATTATCAA	AAAGGATCTT	CACCTAGATC	CTTTTAAATT	2340
AAAAATGAAG	TTTTAAATCA	ATCTAAAGTA	TATATGAGTA	AACTTGGTCT	GACAGTTACC	2400
AATGCTTAAT	CAGTGAGGCA	CCTATCTCAG	CGATCTGTCT	ATTTGCTTCA	TCCATAGTTG	2460
CCTGACTCCC	CGTCGTGTAG	ATAACTACGA	TACGGGAGGG	CTTACCATCT	GGCCCCAGTG	2520
CTGCAATGAT	ACCGCGAGAC	CCACGCTCAC	CGGCTCCAGA	TTTATCAGCA	ATAAACCAGC	2580
CAGCCGGAAG	GGCCGAGCGC	AGAAGTGGTC	CTGCAACTTT	ATCCGCCTCC	ATCCAGTCTA	2640
TTAATTGTTG	CCGGGAAGCT	AGAGTAAGTA	GTTCCGCCAGT	TAATAGTTTG	CGCAACGTTG	2700
TTGCCATTGC	TACAGGCATC	GTGGTGTCAC	GCTCGTCGTT	TGGTATGGCT	TCATTAGCT	2760
CCGGTTCCCA	ACGATCAAGG	CGAGTTACAT	GATCCCCCAT	GTTGTGCAAA	AAAGCGGTTA	2820
GCTCCTTCGG	TCCTCCGATC	GTTGTCAGAA	GTAAGTTGGC	CGCAGTGTTA	TCACTCATGG	2880
TTATGGCAGC	ACTGCATAAT	TCTCTTACTG	TCATGCCATC	CGTAAGATGC	TTTTCTGTGA	2940
CTGTGTAGTA	CTCAACCAAG	TCATTCTGAG	AATAGTGTAT	GCGGCGACCG	AGTTGCTCTT	3000
GCCCGGCGTC	AATACCGGAT	AATACCGCGC	CACATAGCAG	AACTTTAAAA	GTGCTCATCA	3060
TTGGAACACG	TTCTTCGGGG	CGAAAACTCT	CAAGGATCTT	ACCGCTGTTG	AGATCCAGTT	3120
CGATGTAACC	CACTCGTGCA	CCCACTGAT	CTTCAGCATC	TTTTACTTTC	ACCAGCGTTT	3180
CTGGGTGAGC	AAAAACAGGA	AGGCAAAATG	CCGCAAAAAA	GGGAATAAGG	GCGACACGGA	3240
AATGTTGAAT	ACTCATACTC	TTCTTTTTC	AATATTATTG	AAGCATTTAT	CAGGGTTATT	3300
GTCTCATGAG	CGGATACATA	TTTGAATGTA	TTTAGAAAAA	TAAACAAATA	GGGGTTCCGC	3360
GCACATTTCC	CCGAAAAGTG	CCACCTGACG	TCTAAGAAAC	CATTATTATC	ATGACATTAA	3420
CCTATAAAAA	TAGGCGTATC	ACGAGGCCCT	TTCGTCTTCA	AGCTGCCTCG	CGCGTTTCGG	3480
TGATGACGGT	GAAAACCTCT	GACACATGCA	GCTCCCGGAG	ACGGTCACAG	CTTGTCTGTA	3540
AGCGGATGCC	GGGAGCAGAC	AAGCCCGTCA	GGGCGCGTCA	GCGGGTGTTG	GCGGGTGTCG	3600
GGGCGCAGCC	ATGACCCAGT	CACGTAGCGA	TAGCGGAGTT	GGCTTAACTA	TGCGGCATCA	3660
GAGCAGATTG	TACTGAGAGT	GCACCATATC	GACGCTCTCC	CTTATGCGAC	TCCTGCATTA	3720
GGAAGCAGCC	CAGTAGTAGG	TTGAGGCCGT	TGAGCACCGC	CGCCGCAAGG	AATGGTGCTG	3780
GCTTATCGAA	ATTAATCGAC	TCACTATAGG	GAGACCCGAA	TTGAGCTCG	CCCCGTTACA	3840
TAACTTACGG	TAAATGGCCC	GCCTGGCTGA	CCGCCCAACG	ACCCCGCCCC	ATTGACGTCA	3900
ATAATGACGT	ATGTTCCCAT	AGTAACGCCA	ATAGGGACTT	TCCATTGACG	TCAATGGGTG	3960
GAGTATTTAC	GGTAAACTGC	CCACTTGGCA	GTACATCAAG	TGTATCATAT	GCCAAGTACG	4020
CCCCCTATTG	ACGTCAATGA	CGGTAAATGG	CCCGCCTGGC	ATTATGCCCA	GTACATGACC	4080
TTATGGGACT	TTCTTACTTG	GCAGTACATC	TACGTATTAG	TCATCGCTAT	TACCATGGTG	4140
ATGCGGTTTT	GGCAGTACAT	CAATGGGCGT	GGTAGCGGT	TTGACTCACG	GGGATTTCCA	4200
AGTCTCCACC	CAATTGACGT	CAATGGGAGT	TTGTTTTGGC	ACCAAAATCA	ACGGGACTTT	4260

CCAAAATGTC	GTAACAACTC	CGCCCCATTG	ACGCAAATGG	GCGGTAGGCG	TGTACGGTGG	4320
GAGGTCTATA	TAAGCAGAGC	TCGTTTAGTG	AACCGTCAGA	TCGCCTGGAG	ACGCCATCCA	4380
CGCTGTTTTG	ACCTCCATAG	AAGACACCGG	GACCGATCCA	GCCTCCGCGG	GATCTTGGTG	4440
GCGTGAAACT	CCCGCACCTC	TTCGGCCAGC	GCCTTGTAGA	AGCGCGTATG	GCTTCGTGGG	4500
GATCCCCCAA	AGAATCCTTA	GCTCCCCCTG	GTAGAGACGA	AGTCCCTGGC	AGTTTGCTTG	4560
GCCAAGGGAG	GGGGAGCGTA	ATGGACTTTT	ATAAAAGCCT	GAGGGGAGGA	GCTACAGTCA	4620
AGGTTTCTGC	ATCTTCGCCC	TCAGTGGCTG	CTGCTTCTCA	GGCAGATTCC	AAGCAGCAGA	4680
GGATTCTCCT	TGATTTCTCG	AAAGGCTCCA	CAAGCAATGT	GCAGCAGCGA	CAGCAGCAGC	4740
AGCAGCAGCA	GCAGCAGCAG	CAGCAGCAGC	AGCAGCAGCA	GCAGCAGCCA	GGCTTATCCA	4800
AAGCCGTTTC	ACTGTCCATG	GGGCTGTATA	TGGGAGAGAC	AGAAACAAAA	GTGATGGGGA	4860
ATGACTTGGG	CTACCCACAG	CAGGGCCAAC	TTGGCCTTTC	CTCTGGGGAA	ACAGACTTTC	4920
GGCTTCTGGA	AGAAAGCATT	GCAAACCTCA	ATAGTTCGAC	CAGCGTTCCA	GAGAACCCCA	4980
AGAGTTCAAC	GTCTGCAACT	GGGTGTGCTA	CCCCGACAGA	GAAGGAGTTT	CCCCAAACTC	5040
ACTCGGATGC	ATCTTCAGAA	CAGCAAATC	GAAAAAGCCA	GACCGGCACC	AACGGAGGCA	5100
GTGTGAAATT	GTATCCACA	GACCAAAGCA	CCTTTGACCT	CTTGAAGGAT	TTGGAGTTTT	5160
CCGTGGGTC	CCCAAGTAAA	GACACAAACG	AGAGTCCCTG	GAGATCAGAT	CTGTTGATAG	5220
ATGAAAACCT	GCTTTCCTCT	TTGGCGGGAG	AAGATGATCC	ATTCTTCTC	GAAGGGAACA	5280
CGAATGAGGA	TTGTAAGCCT	CTTATTTTAC	CGGACACTAA	ACCTAAAATT	AAGGATACTG	5340
GAGATACAAT	CTTATCAAGT	CCCAGCAGTG	TGGCACTACC	CCAAGTGAAA	ACAGAAAAAG	5400
ATGATTTTCAT	TGAACTTTGC	ACCCCCGGGG	TAATTAAGCA	AGAGAACTG	GGCCCAGTTT	5460
ATTGTCAGGC	AAGCTTTTCT	GGGACAAATA	TAATTGGTAA	TAAATGTCT	GCCATTTCTG	5520
TTCATGGTGT	GAGTACCTCT	GGAGGACAGA	TGTACCACTA	TGACATGAAT	ACAGCATCCC	5580
TTTCTCAGCA	GCAGGATCAG	AAGCCTGTTT	TTAATGTCAT	TCCACCAATT	CCTGTTGGTT	5640
CTGAAAACCTG	GAATAGGTGC	CAAGGCTCCG	GAGAGGACAG	CCTGACTTCC	TTGGGGGCTC	5700
TGAACTTCCC	AGGCCGCTCA	GTGTTTCTA	ATGGGTACTC	AAGCCCTGGA	ATGAGACCAG	5760
ATGTAAGCTC	TCCTCCATCC	AGCTCGTCAG	CAGCCACGGG	ACCACCTCCC	AAGCTCTGCC	5820
TGGTGTGCTC	CGATGAAGCT	TCAGGATGTC	ATTACGGGGT	GCTGACATGT	GGAAGCTGCA	5880
AAGTATTCTT	TAAAAGAGCA	GTGGAAGGAC	AGCACAAATTA	CCTTTGTGCT	GGAAGAAACG	5940
ATTGCATCAT	TGATAAAATT	CGAAGGAAAA	ACTGCCCAGC	ATGCCGCTAT	CGGAAATGTC	6000
TTCAGGCTGG	AATGAACCTT	GAAGCTCGAA	AAACAAAGAA	AAAAATCAAA	GGGATTTCAGC	6060
AAGCCACTGC	AGGAGTCTCA	CAAGACACTT	CGGAAAATCC	TAACAAAACA	ATAGTTCCTG	6120
CAGCATTACC	ACAGCTCACC	CCTACCTTGG	TGTCACTGCT	GGAGGTGATT	GAACCCG	6177

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH:	98 base pairs
(B) TYPE:	nucleic acid
(C) STRANDEDNESS:	single
(D) TOPOLOGY:	linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

GTACGTTTAA	ACGCGGCGCG	CCGTCGACCT	GCAGAAGCTT	ACTAGTGGTA	CCCCATGGAG	60.
ATCTGGATCC	GAATTCACGC	GTTCTAGATT	AATTAAGC			98

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH:	98 base pairs
(B) TYPE:	nucleic acid
(C) STRANDEDNESS:	single
(D) TOPOLOGY:	linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GGCCGCTTAA TTAATCTAGA ACGCGTGAAT TCGGATCCAG ATCTCCATGG GGTACCACTA	60
GTAAGCTTCT GCAGGTCGAC GGC GCGCCGC GTTTAAAC	98

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH:	51 base pairs
(B) TYPE:	nucleic acid
(C) STRANDEDNESS:	single
(D) TOPOLOGY:	linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

GATCTCGGTC TCCAACAGCA ACAGCAACAG CAACAGCAAC AGGGTCTTCT G	51
--	----

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH:	51 base pairs
(B) TYPE:	nucleic acid
(C) STRANDEDNESS:	single
(D) TOPOLOGY:	linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

GATCCAGAAG ACCCTGTTGC TGTGCTGTT GCTGTTGCTG TTGGAGACCG A	51
---	----

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH:	42 base pairs
(B) TYPE:	nucleic acid
(C) STRANDEDNESS:	single
(D) TOPOLOGY:	linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

AATCCCCGA GCGGCAGCT GAAATCATCA CCAATCAGAT CT 42

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH:	18 base pairs
(B) TYPE:	nucleic acid
(C) STRANDEDNESS:	single
(D) TOPOLOGY:	linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

TATGCCTTAC CATGTGGC 18

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH:	25 base pairs
(B) TYPE:	nucleic acid
(C) STRANDEDNESS:	single
(D) TOPOLOGY:	linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

TTGGTCGACA AGATCATGCA TTATC 25

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH:	28 base pairs
(B) TYPE:	nucleic acid
(C) STRANDEDNESS:	single
(D) TOPOLOGY:	linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

TTGTCGACCC GCAGTACAGA TGAAGTTG 28

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH:	30 base pairs
(B) TYPE:	nucleic acid
(C) STRANDEDNESS:	single
(D) TOPOLOGY:	linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

TTGGTCGACC CAGCAATAAC TTCAGACATC 30

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH:	29 base pairs
(B) TYPE:	nucleic acid
(C) STRANDEDNESS:	single
(D) TOPOLOGY:	linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

CGACAGATCT GGCTCCTGAG CAAAGAGAA 29

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH:	24 base pairs
(B) TYPE:	nucleic acid
(C) STRANDEDNESS:	single
(D) TOPOLOGY:	linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

CCAGGGATCC TCTCCTTGCT GCAA

24

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH:	33 base pairs
(B) TYPE:	nucleic acid
(C) STRANDEDNESS:	single
(D) TOPOLOGY:	linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

TCTAGTCGAC GATGGCTCCT GAGCAAAGAG AAG

33

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH:	27 base pairs
(B) TYPE:	nucleic acid
(C) STRANDEDNESS:	single
(D) TOPOLOGY:	linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

CCAGGGATCC TATCCTTGCT GCAACAG

27